PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
A61K 37/02, G01N 33/53, C07K 13/00, C07H 21/02, 21/04, C12N 5/12, 1/21

A1

(11) International Publication Number: WO 94/13312

(43) International Publication Date: 23 June 1994 (23.06.94)

(21) International Application Number: PCT/US93/11404

(22) International Filing Date: 23 November 1993 (23.11.93)

(30) Priority Data:
07/990,866 15 December 1992 (15.12.92) US

(71) Applicant: THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY [US/US]; 900 Welch Road, Suite 350, Stanford, CA 94303 (US).

(72) Inventors: BUTCHER, Eugene; 230 Corte Madera, Portola Valley, CA 94028 (US). BRISKIN, Michael; 1370 Mills Street, #D, Menlo Park, CA 94025 (US).

(74) Agents: ROWLAND, Bertram, I. et al.; Flehr, Hohbach, Test, Albritton & Herbert, 4 Embarcadero Center, Suite 3400, San Francisco, CA 94111-4187 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: MUCOSAL VASCULAR ADDRESSIN, DNA AND EXPRESSION

(57) Abstract

Mammalian purified nucleic acids and proteins encoding mucosal addressins are provided. The proteins, nucleic acids and fragments thereof serve a variety of purposes in modulating leukocyte mucosal tissue homing interactions. In addition, the protein fragments may serve as antigens to produce antibodies specific for particular domains of the mucosal addressin, while the nucleic acids may be used as probes, for antisense preparation, for introduction into hosts which may be allogeneic or xenogeneic to the DNA source for expression of the protein, and the like.

A0

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MR	Mauritania
MW	Malawi
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SI	Slovenia
SK	Slovakia
SN	Senegal
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

MUCOSAL VASCULAR ADDRESSIN, DNA AND EXPRESSION

INTRODUCTION

Technical Field

The field of this invention is the encoding DNA and expression of mucosal vascular addressin.

Background

Blood is the highway of leukocytes, where the leukocytes can move from various tissues, particularly lymphoid tissues, and be transferred to other sites. Since many of the leukocytes monitor for foreign matter and diseased cells, it is necessary that they circulate in a continuous manner to provide protection from disease. Many leukocytes are also involved in providing factors for growth of other cells and aiding in regenerative processes. It is found, however, that certain subsets of leukocytes appear to be programmed to go to different sites. For example, cells will differentially home to mucosal tissue, peripheral lymph nodes, synovial tissue or skin, depending upon the particular proteins which are present on the cell surface. Thus, there are subsets of lymphocytes and monocytes which will preferentially be directed to particularly types of tissue.

The process of a cell being directed to a particular site is referred to as "homing." What this means is that there is a process which allows the cell programmed for a particular site to be directed to that site in the event of injury or disease, or in some instances under normal physiologic circumstances. Part of this process involves endothelial cells, which are part of blood vessels, including high endothelial venules. These cells, present on venules, respond to signals from diseased or injured tissue by upregulating certain surface membrane proteins. The

surface membrane proteins, referred to as addressins, are capable of specifically binding to particular surface membrane proteins of certain lymphocytes or monocytes, providing for the specificity of the circulating cells. After binding to the endothelial cells, the circulating cells may then be directed to the site of injury or disease by a process referred to as "extravasation."

There is substantial interest in being able to control the homing process. There are instances where one would wish to inhibit transfer of cells to a particular site, such as in the case of autoimmune diseases. Alternatively, there may be substantial interest in directing leukocytes to a particular site or for them to be maintained at a particular site, where it may be desirable to concentrate a population of T-lymphocytes or neutrophils. Also, there may be an interest in directing particular compositions to cells or sites for therapeutic or diagnostic purposes. There is, therefor, substantial interest in identifying proteins associated with the homing process, where these proteins may then aid in inhibiting or directing leukocytes or compositions to a designated site in the body.

Relevant Literature

The mechanism of homing of lymphocytes is described by Butcher et al., *Eur. J. Immunol.* 10, 551-556; Butcher, *Curr. Topics Microbiol. Immunol.* 133, 2961-2965 (1986); Picker and Butcher, *Annual Rev. Immunol.* 10, 561-591 (1992); and Springer, *Nature* 346, 425-434 (1990). See also, Streeter et al., *Nature* 331, 41-46 (1988) and Nakachli et al., *Nature* 337, 179-181 (1989). The murine brain endothelioma bEnd3 is described by Montesano et al., *Cell* 62, 435-445 (1990). Descriptions of homing receptor and addressin interactions may be found in Berg

to leukocytes, and the like. Also as a tool for screening to identify molecules that bind to the mucosal addressins, that might themselves find therapeutic uses in inhibiting MAd functions, or might be used to target other agents to MAdCAM-1 and endothelial cells ("EC") for therapeutic purposes. Methods are provided for isolating or synthesizing DNA which encodes at least a portion of the mucosal addressin, for introducing the DNA into host cells, and for expression of the addressin or fragments thereof. In addition, the addressin can serve as a source of the carbohydrate side chains which serve as binding entities to the leukocytes.

BRIEF DESCRIPTION OF THE FIGURES

The Figure depicts the cDNA sequence of murine MAdCAM-1 (Clone 7) and the encoded amino acid sequence. The signal peptide, transmembrane region and single amino acid change (R to L) are bold underlined. In the nucleic acid sequence the base change is T to G (nucleotide 182).

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

Purified nucleic acid and protein compositions are provided, where the protein compositions are mammalian mucosal addressins or fragments thereof and the nucleic acids are sequences encoding the mucosal addressin or fragments thereof. The mucosal addressin protein in the mouse is characterized by binding to antibodies which have been shown to bind to mucosal addressins, MECA-89 and -367. In addition, the mucosal addressin has a structure having three regions, the extracellular region, the transmembrane region, and the cytoplasmic region and when initially expressed, the signal sequence. The mature protein has from about 350-425 amino acids, usually from about 375-400 amino acids. At least about 80% of the mature protein will be the extracytoplasmic portion.

The overall structure of the protein is a Type I transmembrane protein. There are three immunoglobulin-related domains. The most end-terminal domain (domain 1) shows high sequence similarity to domain 1 of the cell adhesion molecule ICAM-1. The second immunoglobulin domain (domain 2) shows sequence similarity to the fifth domain of mouse VCAM-1. Domains 1 and 2 are separated by at least 20 amino acids and not more than about 50 amino acids,

usually about 35-45 amino acids. This is a typical spacing for separation of two adjacent immunoglobulin-like domains.

Just proximal to domain 2 is a long stretch rich in prolines, serines and threonines, indicating a site for heavy O-glycosylation. This domain will be at least about 60 amino acids and not more than about 100 amino acids, generally being from about 65-80 amino acids. Of this domain, the majority of serines and threonines are found in about a 35-40 amino acid stretch. In this region, from about 30-50% of the amino acids will be serine or threonine, while from about 5-20% of the amino acids will be proline.

Proximal to this region is a third immunoglobulin domain, domain 3. This domain has significant homology with a constant region immunoglobulin domain C2 of the IgA1 heavy chain. It should be noted that IgA1 is associated with mucosal tissue.

The mucosal addressin may come from any species, particularly primate, more particularly human, rodentiae, particularly murine, namely mouse and rat, lagomorpha, ovine, bovine, equine, porcine, feline, canine, guinea pig, etc.

The mouse mucosal addressin is characterized by having a signal sequence of 21 amino acids, an extracytoplasmic region of 342 amino acids, a transmembrane sequence of 22 amino acids and a cytoplasmic tail of 20 amino acids. The extracellular region has three potential N-linked glycosylation sites, which do not appear to be used.

The mucosal addressins have substantial homology to other adhesion molecules which have been previously reported, namely VCAM-1 and ICAM-1. This homology is particularly found at the N-terminal domain.

The subject mucosal addressin will hereinafter be referred to as MAdCAM-1.

The subject protein is selectively expressed on HEV of mucosal lymphoid organ HEV and lamina propria venules. The subject protein binds to the previously-reported monoclonal antibodies MECA-89 and -367 and specifically binds to the mucosa-selective T-cell lymphoma TK1. MAdCAM-1 is a member of the immunoglobulin superfamily.

The gene encoding MAdCAM-1 will generally be joined to other than wild-type sequences. Usually the wild-type DNA will be not more than about 1 cM,

generally not more than about 10 kb. The genomic gene will usually be separated from the intact chromosome and may be manipulated as part f a YAC, where fragments will be from about 100-1000 kbp.

The cDNA which encodes the subject protein in the mouse has between 1400-1500 nucleotides (nt) with a short, 5'-noncoding region, generally from about 20-40 nt and a 3'-noncoding region of from about 150-250 nt. The mouse cDNA is 1436 nt, with a 5'-noncoding region of 30 nt, and a 3'-noncoding region of 191 nt. Included in the 3' region is a polyadenylation site at position 1400. The initiation codon, ATG, is at position 31 of the cDNA. The molecule encoded has 21 amino acids for the signal peptide leaving a mature polypeptide of 384 amino acids.

By manipulation of the coding sequence, variations of the mucosal addressin may be achieved. Of particular interest is removal of the transmembrane sequence, either with or without removing the cytoplasmic portion, whereby a soluble secreted form of the mucosal addressin will be obtained. Excising the transmembrane region may be achieved by *in vitro* mutagenesis, polymerase chain reaction, or other known technique.

Expression of the mucosal addressin in endothelial cells, particularly transformed endothelial cells, can be enhanced by using TNF-α or lipopolysaccharide, with relatively high levels of expression occurring within a few hours, generally within about 4-6 hr.

Expression of the protein in an appropriate mammalian host will provide for the glycosylated mature product, which may be isolated and purified in accordance with conventional ways. The protein can be obtained in a pure state of at least about 50 wt.%, preferably at least about 75% and more preferably substantially pure or pure, in that the composition is free of any physiologically active contaminants.

The various compositions provided by the subject invention may be used in a variety of ways. The intact protein, fragments of the protein indicating polypeptide fragments, polysaccharides, and combinations thereof may find application. The DNA encoding the mucosal addressin, the entire reading frame, with or without introns, non-coding regulation elements of the gene, or fragments of the encoding DNA, particularly fragments encoding all or a portion of a

functional region or domain can find use. The sequences may be mutated for a variety of reasons, where the mutations may include insertions, deletions and substitutions. Substitutions may be conservative or non-conservative, there will usually be fewer than the greater of 2 substitutions or substitutions equal to 2% of the number of nucleotides. For the protein, the same numbers will apply for the amino acids.

Polypeptides of each of the domains may be used to prepare a specific antisera or monoclonal antibodies for a specific site. In this way, binding to each of the domains may be individually inhibited by using antibodies which specifically bind to that domain. These antibodies may then be used to determine the manner in which binding occurs between the addressin and various homing receptors or selectins. The glycosylated domain may be used in a variety of ways, such as providing a source of the sugars which may bind to lectin regions of the selectins, for producing antibodies to the protein and/or saccharide portions of the domain, to serve as inhibitors to binding, to direct compositions to activated leukocytes, and the like.

The protein or fragments thereof, usually fragments of at least about 12 amino acids, more usually at least about 16 amino acids and preferably at least about 20 amino acids, generally fewer than the entire protein, preferably fewer than 60 amino acids may be employed as agents to block binding of the leukocytes to the mucosal addressin. These peptides may find use *in vitro* and *in vivo*, where the peptides may inhibit entry of leukocytes to a site where their presence may contribute to a disease. Thus, one may wish to exclude leukocytes from injured sites, where the leukocytes may enhance inflammation, induce cellular proliferation, destroy native tissue, or the like. Diseases which may be treated include especially intestinal inflammatory disorders including Crohn's disease, ulcerative colitis, regional enteritis, and Celiac disease; but also in inflammatory and autoimmune disorder in other sites where the MAdCAM-1 may be induced abnormally or in pathologic conditions. Also in malignancies, especially of intestinal carcinomas and their metastases, in which MAdCAM-1 may be induced in tumor-associated vessels.

The protein or fragments thereof may be modified in a variety of ways for a variety of purposes. The peptides may be joined to immunogens to enhance

immunogenicity for the formation of antibodies. The antibodies may be from any convenient source but will normally be produced by laboratory animals, such as mice, rats and rabbits, or in domestic animals, such as cows, sheep, pigs, etc. The manner of producing antisera and monoclonal antibodies is sufficiently well-established as to not require description here. The mouse or other species monoclonal antibodies may be humanized by replacing the constant region, by itself or in combination with one or more conserved framework regions, with regions from human antibodies, so that human IgA, -G, or -M or the like regions may replace the native antibody regions to reduce antigenicity in humans and to provide effector functions.

The protein or peptides may be joined to polyethylene glycol, constant regions of immunoglobulins, e.g. IgG, to lipids or the like, to enhance the lifetime of the peptide or protein when administered *in vivo*. The peptide or protein may be joined to various effector agents, particularly agents which produce a physiological reaction, such as antibodies, toxins, drugs, carbohydrates or cytokines. The peptide or protein may be joined to these various agents in a variety of ways. Where proteins are involved, the sequence encoding the subject peptide or protein may be fused to the other protein to provide for a fusion protein. Where a lipid is to be joined, it may be feasible to provide for the coding sequence which encodes for the linking of the peptide to a lipid through a phosphate linkage, based on known, natural processes of the host cell. Alternatively, various chemical linking agents are known which allow for linking through carboxyl groups, amino groups, thiol groups, or hydroxyl groups, as well as linking through an activated aromatic group, such as is present in tyrosine and tryptophan. A description of these agents may be found in *Antibodies, A Laboratory Manual*, Harlow, E., Lane, D. ed., Cold Spring Harbor Laboratory, 1988, pp. 319-358. Various synthetic procedures and conditions are known, which need not be described here.

In addition, the peptides or proteins may be labeled with a wide variety of labels for a variety of purposes. Thus, the peptides or proteins may be used in assays for detecting the presence of cells binding to the subject mucosal addressin or fragment thereof. Thus, fluorescent labels, enzyme labels, radioactive labels, or the like may be employed for use of the subject peptides or proteins in the assays.

Numerous assays are known, such as RIA, ELISA, fluorescence polarization assay, fluorescence-activated cell sorters (FACS), and the like.

In addition, the peptides or proteins may be bound to a solid support, such as particles, microtiter plate wells, solid surfaces, where the peptides may serve to bind to cells having the appropriate complementary protein. In this manner, cells may be selectively removed *ex vivo* for isolation and characterization, to reduce their population in a host, or the like. The particles may be magnetic particles, pore glass beads, latex particles, agarose particles, sepharose particles, or the like. The various particles may be activated in a wide variety of ways, activated particles being commercially available for covalently linking the peptide or protein to the particle. Similarly, other surfaces may also be activated by known ways for covalent binding. Alternatively, depending upon the purpose of the binding, with some surfaces, e.g. polystyrene, non-covalent binding may suffice. After binding of the subject peptide or protein, usually an innocuous protein will be used to react with any remaining unreacted functional groups and to coat any hot spots which might result in non-specific binding in a subsequent application.

The DNA encoding the subject addressin or fragment thereof may be used for a diversity of purposes. The cDNA or fragment thereof may be used to identify the genomic gene which encodes the subject addressin. A cDNA or fragment thereof from one species may be used to probe cDNA or genomic libraries of other species to identify and isolate the gene from the other species. In addition, one may use the sequence encoding the mucosal addressin to define antisense sequences, which may then be used to inhibit expression of the gene. In this way, one can inhibit the expression of the mucosal addressin in mucosal endothelial cells, thus inhibiting the homing of lymphocytes via these endothelial cells.

One may also use the sequences for homologous recombination, so that the sequences may be used to correct a defective mucosal addressin gene, to specifically knock-out a mucosal addressin gene or to provide for enhanced or regulated expression of the mucosal addressin gene, by changing the promoter or other regulatory sequence present upstream from the coding region. In this way, one can investigate the effect on endothelial cells of the inability to express the mucosal addressin.

The subject peptides may be used in assays to identify drugs, including peptides, saccharides and small synthetic organic molecules which bind to the peptides and have biological activity. These agents may then be used in assays to determine their effectiveness *in vivo*. The initial assays may take various forms involving direct binding or competitive binding. In the former, the subject peptides may be bound to a surface and the radiolabeled agent added. By determining the radioactivity bound to the surface after washing to remove non-specific binding, the affinity of the agent may be determined. Alternatively, one can add the agent to a complex of a known binding agent and the peptide and determine the off rate or equilibrium concentration of the labeled known binding agent as an indication of the affinity of the candidate agent.

The proteins or fragments thereof, capable of binding to the complementary binding protein may also be used as antagonists for complex formation and to modulate the various interactions. Thus, by administering the respective protein or fragment thereof to a host, the protein may serve to home to the complementary binding member and inhibit the binding of the addressin associated with the target cells. The fragment will usually be at least 8 amino acids, more usually at least about 10 amino acids, frequently at least about 16 amino acids, usually not exceeding about 60 amino acids, preferably not exceeding about 36 amino acids.

Rather than acting as inhibitors or enhancers to prevent or increase complex formation between leukocytes and target sites, the proteins, fragments thereof, or anti-idiotope antibodies may serve to direct a wide variety of molecules to the target site. Thus, in the case of neoplastic tissue, by administering one of the subject compounds or compositions bound to a therapeutic drug, one can direct the binding of a therapeutic drug to the desired site for retention, and concentration at the desired site to which the MAdCAM-1 expressing endothelial cells bind. One can provide for the binding of radioisotopes for *in vivo* diagnosis or imaging, for radiotherapy, or the like. Alternatively, one could bind cytotoxic drugs to the fragments to kill target lymphocytes.

In administering the various therapeutic agents, for the most part, empirical determinations will be involved to determine the level of therapeutic agent. The level of therapeutic agent which is administered will depend to a substantial degree

administration, the purpose of the therapy, and the like. Therefore, no simple range may be given which would indicate what levels should be applied for any particular therapy. For the most part, the proteins will be administered in an appropriate physiologically-acceptable medium, e.g. water, saline, phosphate-buffered saline, mineral oil, Ringer's solution, or the like. Administration will normally be parenteral, particular intravascularly, but may be oral. For the reasons given above, the course of treatment will also vary.

There may be situations where there is an interest in using cells as therapeutics. By employing the subject peptides, one may isolate particular subsets of cells, which may then be expanded in culture in an appropriate nutrient medium or in co-culture, and may then be used for a therapeutic purpose. Depending upon the particular situation, the cells may be autologous or allogeneic.

The antibodies may also find wide use for detecting the presence of venules *in vitro* or *in vivo*, where the subject mucosal addressin is upregulated. By using antibodies, binding of leukocytes to these sites may be inhibited, so as to prevent extravasation to the particular site. The antibodies may also be used to make anti-id antibodies, which may be selected as mimics for the subject peptides or proteins. In this way, the antibodies which will have a substantially different composition from the mucosal addressin, may fulfill the same function by binding to the endothelial cell and block binding of the leukocyte to the venule. Also, the antibodies may be used in the variety of ways indicated above in place of the subject peptides or proteins.

The cDNA or genomic DNA encoding MAdCAM-1 can be used in preparing expression constructs. The DNA can be used in any convenient host, both prokaryotic and eukaryotic for the cDNA and invertebrate and vertebrate for splicing capability for the genomic gene. As previously indicated, for glycosylation, eukaryotic, particularly mammalian hosts will be required. The gene encoding MAdCAM-1 may be inserted into a wide variety of generally-available expression cassettes having linkers between a 5' transcriptional and translational initiation region, which may include an enhancing sequence, and may provide for constitutive or inducible transcription, and a 3' transcriptional and translational termination region, which will normally include a polyadenylation signal. Alternatively, one may prepare an expression cassette by combining the

appropriate functional sequences in the appropriate order as indicated above. Numerous transcriptional initiation regions are commercially available or described in the literature. The expression cassette may then be introduced into a vector for replication and expansion. The resulting vector will usually include a marker for selection of hosts containing the construct. The vector may be introduced into the expression host by any convenient means, precipitated DNA, transfection, fusion, electroporation, etc., where the vector may provide for extrachromosomal maintenance or integration into the genome of the host. Of particular interest are vectors comprising sequences of a virus for transfection and integration into mammalian cellular hosts. The expression host may then be grown. Depending on the host, the product may be secreted and isolated from the supernatant or in the event it is retained in the cytoplasm, the cells may be harvested and the lysate extracted. The product may be purified using electrophoresis, gel chromatography, affinity chromatography, extraction and the like.

The following examples are offered by way illustration and not by way limitation.

EXPERIMENTAL

A transformed mouse endothelioma cell line, Bend3 cells was employed. It was shown to be inducible for the MECA-89 and -367 epitopes with bacterial lipopolysaccharide (LPS) and tumor necrosis factor α (TNF-α). Serial fluorescence activated cell sorting was used to generate a variant line termed Hi-MAd 4 that expresses much higher levels of the antigen upon TNF-α stimulation. Hi-MAd 4 cells, stimulated with recombinant TNF-α for either 6 or 8 hours were pooled and used as the source of RNA for cDNA libraries. Endothelial cell cDNA libraries were synthesized with a BRL system which generates directionally-oriented cDNA into the λgt11 derivative, λgt22A. Phage clones (5×10^5) were screened with a mixture of anti-mucosal addressin antibodies consisting of a rat polyclonal against affinity-purified MAd (1:2000) and the monoclonal antibodies MECA-89 and -367 (5 μg/ml) diluted TBS with 1% BSA and 0.02% sodium azide. Positive clones were detected with an anti-rat IgG conjugated to alkaline phosphatase (Promega) and BioRad BCIP/NBT color

The entire sequence of the mucosal addressin cDNA is 1436 nt in length, consisting of a 30 nt 5' untranslated region and a 191 nt 3' untranslated region. A canonical polyadenylation signal AATAAA is located at position 1400. The longest open reading frame of 405 amino acids begins with the first ATG at position 31 and ends with the termination codon TGA at position 1245. Hydrophobicity analysis and rules governing signal peptide cleavage patterns predict a signal peptide of 21 amino acids, leaving a mature polypeptide of 384 amino acids. This sequence predicts a typical transmembrane protein consisting of a 342 amino acid extracellular domain, a 22 amino acid transmembrane region and a short, charged 20 amino acid cytoplasmic tail.

The extracellular domain contains 3 potential N-linked glycosylation sites with the consensus sequence NXT/S at position 718, 787 and 111. Studies indicate that these sites are not utilized, since the mucosal addressin is insensitive to peptide N-glycosidase.

The largest cDNA inserts from λ clones 7 and 15 (λ MAd-7 and λ MAd-15) were subcloned into Bluescript vectors (Stratagene) to construct the plasmids pMAd-7 and pMAd-15. Plasmids for sequencing were initially generated by making nested deletions with exonuclease III using Erase A Base (Promega) system. The DNA sequence was determined by the dideoxy method using fluorescent-labeled universal primers and the Applied Biosystems DNA sequence analyzer. Gaps in the sequence were filled in by subcloning small restriction fragments into appropriately-digested Bluescript vectors. Two independent clones were sequenced 100% on both strands and are identical with the exception of a single base substitution which would generate a leucine instead of the arginine shown in pMAd-15 at amino acid 61.

In the sequence provided in the Figure, the putative hydrophobic signal peptide and transmembrane sequences have a bold underline. Potential N-linked glycosylation sites are present at positions 718, 787 and 111. Cysteines likely to form disulfide bonds in Ig domains are boxed while other cysteines are underlined. The 37 amino acid "mucin domain" between amino acid residues 221 and 257 is outlined with a bold line. The polyadenylation signal ATAAA at position 1400 is indicated in bold.

Poly A+ RNA from unstimulated high MAd4 cells, high MAd-4 cells stimulated with TNF-α for 6-8 hrs, high MAd-4 cells stimulated with TNF-α for 20 hrs, the L1-2 pre-B lymphoma, normal 8 wk. BALB/c mouse liver, kidney, Peyer's patches, brain, peripheral lymph nodes, lung, mesenteric lymph nodes and spleen were hybridized to the 1.4 kb MAd cDNA insert. Northern blots used 1 μg of poly A+ RNA from cells lines and 2 μg from tissues which was denatured and electrophoresed through a 1% agarose formaldehyde gel and transferred to a PVDF (Immobilon, Millipore) membrane by standard capillary blot procedures. Equal amounts (by OD_{260}) of each RNA were electrophoresed and stained with ethidium bromide to ensure equivalent loading of each sample. Hybridization was performed at 65°C using standard conditions. cDNAs were labeled with $\alpha^{32}P$ dCTP by priming with random hexamers.

Northern blot analysis showed that two species of RNA at approximately 1.8 and 1.3 kb were slightly induced in high MAd-4 cells stimulated with TNF-α for 6-8 hrs and were highly induced by 20 hrs. The same transcripts are highly expressed in mesenteric lymph nodes (MLN) and Peyer's patches (PP) and detected at lower levels in spleen and peripheral lymph nodes (PLN). The species were not detected with the pre-B lymphoma L1-2 or in liver, brain, kidney or thymus. This pattern of expression is consistent with previous immunohistological studies which have shown high expression of MAd in HEV, in PP, mesenteric lymph node (MLN) and some expression in the marginal sinus around splenic white pulp nodules in the spleen. A distinct transcript of approximately 1.5 kb was detected in unstimulated high MAd-4 cells, L1-2 cells, kidney and liver, implying either a related gene or alternatively spliced product in these tissues.

The expression construct pCDM15 was generated by subcloning the Sal1-Not1 insert of λMAd-15 into XhoI-NotI digested pCDM8. This construct was transfected onto COS cell monolayers (along with no DNA as a mock transfected control) using DEAE dextran. After 48 hrs., cells were replated into 3.5 cm dishes to ensure equal plating of all transfectants and 24 hrs. later transfectants were analyzed by immunofluorescence. Both addressin and mock transfectants were stained with anti-addressin monoclonals MECA-89 or -367 alone or the isotype matched control MEL-14. Staining was detected both visually and by

FACS analysis using a phycoerythrin-conjugated goat $F(ab')_2$ anti rat IgG (Caltag). Transfection efficiencies were approximately 25% for each experiment.

For adhesion assays, COS cell transfectants were plated into 3.5 cm plates as described above. B and T lymphomas were pelleted and resuspended at a density of 2×10^6 per ml in RPMI without sodium carbonate, with 2% serum and buffered with 10 mM HEPES, pH 6.7. Cell suspensions were layered (1 ml per plate) onto addressin and mock transfectants and gently rotated on a gyrotary shaker (New Brunswick Scientific Model G-2) at 60 rpm for 30 min. at 25°C, and then washed 4 times in the same medium by aspirating and gently filling each plate with 1 ml per wash. Assays were fixed in Dulbecco's PBS containing 1% glutaraldehyde and sodium azide. For antibody inhibition assays, transfectants were pre-treated with 20 μg/ml of anti-addressin antibody MECA-367 or isotype matched control 9B5 (anti-CD44) at 1 ml per plate for 20 min. at 25°C. Solutions were then aspirated and adhesion assays were performed as described above. The average number of cells bound for each experiment was determined by counting cells in 4 random fields.

Both immunofluorescence and FACS analysis showed that COS cells transfected with pCDM15 (the full-length addressin cDNA inserted into the CDM8 expression vector) expressed antigen which reacts with anti-mucosal addressin MAbs MECA-89 and -367, but not with isotype matched control MAbs. To assay the function of the cloned MAd cDNA, adhesion of the mucosal HEV binding lymphoma cell line TK1 to the COS cell transfectants showed that TK1 cells bind to pCDM15 transfected COS cells, but not to control (mock) transfectants. Binding is specific to cells expressing MAd as shown by parallel immunofluorescence with a non-blocking MAd, MECA-89. Adhesion is inhibited by approximately 95% by pre-treatment with saturating levels of anti-addressin MAd MECA-367, while pre-treatment with control MAd had no effect. Furthermore, binding was specific for a mucosal HEV binding cell line, as L1-2 and NS8 cells (lines previously shown not to adhere to HEV) failed to adhere to the COS cell transfectants.

The amino acid sequence of the mucosal addressin was analyzed for homology to other proteins in the NBRF database using the program FASTP and the highest degree of homology was to murine ICAM-1. Consistent with the

designation f ICAM-1 as a member of the immunoglobulin supergene family, the extracellular domain of MAd can also be folded into a structure containing 3 immunoglobulin-like domains. Each of the 3 domains contains an invariant pair of cysteine residues known to stabilize the Ig loop, with distances of 44, 66 and 47 amino acids between cysteines, respectively.

Additional searches with the individual immunoglobulin-like domains of MAd revealed the highest homology of the most N-terminal domain (domain 1) with the first domain of rat ICAM-1 and the first domain of human VCAM-1. The first domain in all three molecules is similar in that the spacing between the cysteine residues and the predicted β-strand structure is typical of H or C2 immunoglobulin domains. Additionally, all three molecules have the unusual feature of double cysteine residues at both ends of the domain. The N-terminal Ig domains of ICAM-1 and VCAM-1 have been shown to contain binding sites for their respective lymphocyte counter-receptors LFA-1/$\alpha_L\beta_2$ and VLA-4/$\alpha_4\beta_1$. The $\alpha_4\beta_7$ integrin is known to be involved in lymphocyte binding to Peyer's patch-HEV in *in vitro* assays. The present finding of strong homology to other immunoglobulin-related vascular integrin receptors, supports an integrin binding motif lying in this domain in MAdCAM-1.

The second immunoglobulin domain is most closely related to the fifth domain of mouse VCAM-1. The third immunoglobulin-like domain of the MAd is unique in that the most significant homology is to the CH2 region of human and gorilla IgA1 (33% identity). This homology relates two molecules which are highly expressed in mucosal tissues and which both interact with T-cells involved in mucosal immunity. A discrete subset of T-cells, enriched in Peyer's patches, constitutively expressed Fc receptors capable of binding IgA. Their binding to mucosal addressin could be strengthened by interaction of the IgA F_c receptor or other T-cell surface receptors, which are capable of binding the IgA-1 homologous domain.

A proline/serine-threonine-rich region in the amino-terminal portion of this domain which aligns with the hinge region encoded with the IgA1 CH2 exon may add flexibility to the addressin and its interaction with ligands on the surface of lymphocytes.

Th 37 amino acid region that lies between d mains 2 and 3, which is extremely rich in its serine/threonine content (41%), exhibits the sequence characteristic of mucins. This structure is characteristic of O-glycosylation sites. These domains are suggested to form rigid rods without secondary structure and may serve to extend ligand binding domains further above the cell surface, thus allowing more efficient interactions. Rosen and colleagues (Lasky et al., *Cell* 69, 927-938 (1992)) have described a 50 kb mucin-like protein GlyCAM-1, which is highly expressed in peripheral and mesenteric lymph nodes. The protein core of GlyCAM-1 is also serine/threonine rich and is proposed to serve as a backbone molecule capable of presenting a carbohydrate ligand(s) characterizing the peripheral lymph node addressin. These carbohydrate structures are recognized by the L-selectin lymph node homing receptor and by anti-PLN addressin MAb MECA-79. As MECA-79-defined epitopes associated with the peripheral lymph node addressin also decorate a portion of MAdCAM-1 expressed in mesenteric lymph nodes, this mucin-like domain in the mucosal addressin can serve as a site for addition of selectin-binding carbohydrates important for recruitment of lymphocyte subsets to the mesenteric lymph node and also, at a lower level, to the Peyer's patches. Proteoglycans can also bind and present heparin-binding cytokines at the cell surface. Thus, when appropriately modified, the mucin-like region may also be available to bind adhesion-promoting cytokine-activating factors required for triggering activation-dependent lymphocyte sticking to the HEV.

It is evident from the above results, that the subject invention provides an opportunity for modulating homing to mucosal tissue by a great diversity of compositions and protocols. In addition, the compositions can be used for diagnostic and therapeutic purposes, for isolation of subsets of lymphocytes or other leukocytes, and the like.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this

invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: THE BOARD OF TRUSTEES OF LELAND STANFORD JUNIOR UNIVERSITY

(ii) TITLE OF INVENTION: MUCOSAL VASCULAR ADDRESSIN, DNA AND EXPRESSION

(iii) NUMBER OF SEQUENCES: 2

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Bertram I. Rowland
(B) STREET: 4 Embarcadero Center
(C) CITY: San Francisco
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94111

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/US93/
(B) FILING DATE: 23-Nov-1993
(C) CLASSIFICATION: 530

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Rowland, Bertram I.
(B) REGISTRATION NUMBER: 20,015
(C) REFERENCE/DOCKET NUMBER: FP-57452/BIR;STAN-144

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (415) 781-1989
(B) TELEFAX: (415) 398-3249

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1434 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 31..1245

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CCACGCGTCC GTCACAAGAC AGAGGCAGGC ATG GAA TCC ATC CTG GCC CTC CTG      54
                                 Met Glu Ser Ile Leu Ala Leu Leu
                                   1               5

CTG GCC CTG GCC CTA GTA CCC TAC CAG CTC AGC AGA GGA CAG TCC TTC     102
Leu Ala Leu Ala Leu Val Pro Tyr Gln Leu Ser Arg Gly Gln Ser Phe
     10                  15                  20

CAG GTG AAC CCC CCT GAG TCT GAG GTA GCT GTG GCC ATG GGC ACA TCC     150
Gln Val Asn Pro Pro Glu Ser Glu Val Ala Val Ala Met Gly Thr Ser
 25                  30                  35                  40

CTC CAG ATC ACC TGC AGC ATG TCC TGT GAC GAG GGT GTA GCC CGG GTG     198
Leu Gln Ile Thr Cys Ser Met Ser Cys Asp Glu Gly Val Ala Arg Val
                 45                  50                  55

CAC TGG CGT GGT CTG GAC ACC AGC TTG GGC AGT GTA CAG ACC CTC CCA     246
His Trp Arg Gly Leu Asp Thr Ser Leu Gly Ser Val Gln Thr Leu Pro
             60                  65                  70

GGC AGC AGT ATC CTC TCT GTA CGG GGC ATG CTG TCA GAC ACA GGC ACT     294
Gly Ser Ser Ile Leu Ser Val Arg Gly Met Leu Ser Asp Thr Gly Thr
         75                  80                  85

CCT GTG TGT GTG GGC TCC TGC GGG AGT CGA AGC TTC CAG CAC TCC GTG     342
Pro Val Cys Val Gly Ser Cys Gly Ser Arg Ser Phe Gln His Ser Val
     90                  95                  100

AAG ATC CTT GTG TAT GCC TTC CCA GAC CAG CTG GTG GTG TCC CCG GAG     390
Lys Ile Leu Val Tyr Ala Phe Pro Asp Gln Leu Val Val Ser Pro Glu
105                 110                 115                 120

TTC CTT GTA CCT GGA CAG GAC CAG GTG GTG TCC TGC ACG GCC CAC AAC     438
Phe Leu Val Pro Gly Gln Asp Gln Val Val Ser Cys Thr Ala His Asn
                125                 130                 135

ATC TGG CCT GCA GAC CCG AAC AGT CTC TCC TTT GCC CTG CTA CTG GGA     486
Ile Trp Pro Ala Asp Pro Asn Ser Leu Ser Phe Ala Leu Leu Leu Gly
            140                 145                 150

GAG CAG AGA CTG GAG GGT GCC CAA GCC CTG GAA CCA GAG CAA GAA GAG     534
Glu Gln Arg Leu Glu Gly Ala Gln Ala Leu Glu Pro Glu Gln Glu Glu
        155                 160                 165

GAG ATA CAA GAG GCT GAG GGC ACA CCA CTG TTC CGA ATG ACA CAA CGC     582
Glu Ile Gln Glu Ala Glu Gly Thr Pro Leu Phe Arg Met Thr Gln Arg
    170                 175                 180

TGG CGG TTA CCC TCC CTG GGG ACC CCT GCC CCT CCT GCC CTT CAC TGC     630
Trp Arg Leu Pro Ser Leu Gly Thr Pro Ala Pro Pro Ala Leu His Cys
185                 190                 195                 200

CAG GTC ACC ATG CAG CTG CCC AAA CTG GTG CTG ACC CAT AGA AAG GAG     678
Gln Val Thr Met Gln Leu Pro Lys Leu Val Leu Thr His Arg Lys Glu
                205                 210                 215
```

```
ATT CCA GTA CTA CAG AGC CAG ACC TCA CCT AAG CCC CCC AAC ACG ACC      726
Ile Pro Val Leu Gln Ser Gln Thr Ser Pro Lys Pro Pro Asn Thr Thr
            220                 225                 230

TCT GCT GAG CCC TAC ATC CTG ACC TCA TCA AGT ACT GCT GAG GCA GTC      774
Ser Ala Glu Pro Tyr Ile Leu Thr Ser Ser Ser Thr Ala Glu Ala Val
        235                 240                 245

TCC ACT GGG CTC AAC ATC ACC ACC CTA CCT TCT GCC CCT CCA TAC CCC      822
Ser Thr Gly Leu Asn Ile Thr Thr Leu Pro Ser Ala Pro Pro Tyr Pro
    250                 255                 260

AAG CTT AGC CCT AGG ACT CTG AGC TCT GAG GGA CCT TGC CGC CCG AAA      870
Lys Leu Ser Pro Arg Thr Leu Ser Ser Glu Gly Pro Cys Arg Pro Lys
265                 270                 275                 280

ATC CAC CAG GAC CTG GAG GCA GGC TGG GAG CTA CTC TGT GAA GCA TCC      918
Ile His Gln Asp Leu Glu Ala Gly Trp Glu Leu Leu Cys Glu Ala Ser
                285                 290                 295

TGT GGG CCC GGA GTT ACT GTG CGC TGG ACC TTG GCT CCT GGC GAC CTG      966
Cys Gly Pro Gly Val Thr Val Arg Trp Thr Leu Ala Pro Gly Asp Leu
            300                 305                 310

GCA ACC TAC CAC AAG AGG GAG GCT GGG GCC CAG GCA TGG CTA AGC GTG     1014
Ala Thr Tyr His Lys Arg Glu Ala Gly Ala Gln Ala Trp Leu Ser Val
        315                 320                 325

CTG CCC CCA GGT CCC ATG GTA GAG GGC TGG TTC CAG TGC CGC CAG GAC     1062
Leu Pro Pro Gly Pro Met Val Glu Gly Trp Phe Gln Cys Arg Gln Asp
    330                 335                 340

CCT GGC GGG CAG GTG ACC AAT CTG TAT GTT CCT GGC CAG GTG ACC CCG     1110
Pro Gly Gly Gln Val Thr Asn Leu Tyr Val Pro Gly Gln Val Thr Pro
345                 350                 355                 360

AAT TCC TCC TCC ACC GTC GTC CTA TGG ATT GGC AGC TTG GTG CTG GGG     1158
Asn Ser Ser Ser Thr Val Val Leu Trp Ile Gly Ser Leu Val Leu Gly
                365                 370                 375

CTG CTT GCA CTG GTC TTC CTT GCC TAC CGC CTG TGG AAA TGC TAC CGG     1206
Leu Leu Ala Leu Val Phe Leu Ala Tyr Arg Leu Trp Lys Cys Tyr Arg
            380                 385                 390

CCA GGT CCT CGC CCA GAC ACT AGC TCA TGT ACA CAC CTA TGAAGCTCCA      1255
Pro Gly Pro Arg Pro Asp Thr Ser Ser Cys Thr His Leu
        395                 400                 405

TTATGCCAGA CTAAAGGAGG CAGAGAGTGA CCAGCTGCAG GATTTGGGGC ATCAAGATGA  1315

TAGTGTGGCC TCTTTCCTTG GTGGTCAGCA CATCTATAAG TTTCTCCTGA CTTCTGGGCT  1375

TTTCTGCCTG CTGGCCCAGA GCTAAATAAA AGCCCCGTAT CTAAAAAAAA AAAAAAAAA   1434
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 405 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu Ser Ile Leu Ala Leu Leu Leu Ala Leu Ala Leu Val Pro Tyr
1 5 10 15

Gln Leu Ser Arg Gly Gln Ser Phe Gln Val Asn Pro Pro Glu Ser Glu
20 25 30

Val Ala Val Ala Met Gly Thr Ser Leu Gln Ile Thr Cys Ser Met Ser
35 40 45

Cys Asp Glu Gly Val Ala Arg Val His Trp Arg Gly Leu Asp Thr Ser
50 55 60

Leu Gly Ser Val Gln Thr Leu Pro Gly Ser Ser Ile Leu Ser Val Arg
65 70 75 80

Gly Met Leu Ser Asp Thr Gly Thr Pro Val Cys Val Gly Ser Cys Gly
85 90 95

Ser Arg Ser Phe Gln His Ser Val Lys Ile Leu Val Tyr Ala Phe Pro
100 105 110

Asp Gln Leu Val Val Ser Pro Glu Phe Leu Val Pro Gly Gln Asp Gln
115 120 125

Val Val Ser Cys Thr Ala His Asn Ile Trp Pro Ala Asp Pro Asn Ser
130 135 140

Leu Ser Phe Ala Leu Leu Leu Gly Glu Gln Arg Leu Glu Gly Ala Gln
145 150 155 160

Ala Leu Glu Pro Glu Gln Glu Glu Glu Ile Gln Glu Ala Glu Gly Thr
165 170 175

Pro Leu Phe Arg Met Thr Gln Arg Trp Arg Leu Pro Ser Leu Gly Thr
180 185 190

Pro Ala Pro Pro Ala Leu His Cys Gln Val Thr Met Gln Leu Pro Lys
195 200 205

Leu Val Leu Thr His Arg Lys Glu Ile Pro Val Leu Gln Ser Gln Thr
210 215 220

Ser Pro Lys Pro Pro Asn Thr Thr Ser Ala Glu Pro Tyr Ile Leu Thr
225 230 235 240

Ser Ser Ser Thr Ala Glu Ala Val Ser Thr Gly Leu Asn Ile Thr Thr
245 250 255

Leu Pro Ser Ala Pro Pro Tyr Pro Lys Leu Ser Pro Arg Thr Leu Ser
260 265 270

Ser Glu Gly Pro Cys Arg Pro Lys Ile His Gln Asp Leu Glu Ala Gly
275 280 285

Trp Glu Leu Leu Cys Glu Ala Ser Cys Gly Pro Gly Val Thr Val Arg
290 295 300

Trp Thr Leu Ala Pro Gly Asp Leu Ala Thr Tyr His Lys Arg Glu Ala
305 310 315 320

Gly Ala Gln Ala Trp Leu Ser Val Leu Pro Pro Gly Pro Met Val Glu
325 330 335

Gly Trp Phe Gln Cys Arg Gln Asp Pro Gly Gly Gln Val Thr Asn Leu
340 345 350

Tyr Val Pro Gly Gln Val Thr Pro Asn Ser Ser Ser Thr Val Val Leu
355 360 365

Trp Ile Gly Ser Leu Val Leu Gly Leu Leu Ala Leu Val Phe Leu Ala
370 375 380

Tyr Arg Leu Trp Lys Cys Tyr Arg Pro Gly Pro Arg Pro Asp Thr Ser
385 390 395 400

Ser Cys Thr His Leu
405

WHAT IS CLAIMED IS:

1. An isolated DNA sequence other than as a part of a chromosome encoding a mucosal addressin characterized by having three immunoglobulin-like domains, a mucin-like region, between the second and third domains, a transmembrane domain and a cytoplasmic domain, wherein said mucosal addressin is capable of binding to a leukocyte surface membrane protein, or single or double stranded fragment of said DNA of at least about 12 nt in length.

2. An isolated DNA according to Claim 1, wherein said DNA is cDNA.

3. An isolated DNA according to Claim 1, wherein said cDNA is mouse cDNA.

4. An isolated DNA according to Claim 1, wherein said fragment encodes at least one of said immunoglobulin like domains or said mucin-like domain.

5. An isolated DNA according to Claim 1, wherein said fragment encodes cDNA of a differentially spliced mucosal addressin.

6. An isolated DNA according to Claim 1, wherein said fragment encodes a soluble form of said mucosal addressin including at least the extracellular portion of said mucosal addressin.

7. Purified MAdCAM-1 characterized by having three immunoglobulin-like domains, a mucin-like region, between the second and third domains, a transmembrane domain and a cytoplasmic domain and being capable of binding to a leukocyte surface membrane protein or fragment thereof of at least about 12 amino acids.

8. Purified MAdCAM-1 according to Claim 7, wherein said MAdCAM-1 is mouse MAdCAM-1.

9. Purified MAdCAM-1 according to Claim 7, wherein said MAdCAM-1 comprises a fragment of MAdCAM-1 comprising at least one immunoglobulin domain or said mucin region.

10. Purified MAdCAM-1 according to Claim 7, wherein said MAdCAM-1 is the expression product of a genetic construct in an expression host.

11. Purified MAdCAM-1 according to Claim 10, wherein said host is a prokaryotic host and said MAdCAM-1 lacks glycosylation.

12. Purified MAdCAM-1 according to Claim 10, wherein said host is a eukaryotic host.

13. A purified fragment of MAdCAM-1, comprising the extracellular portion of MAdCAM-1, said fragment being soluble in blood.

14. A genetic construct comprising a DNA sequence according to Claim 1 under the transcriptional and translational control of transcriptional and translational initiation and termination regions functional in an expression host.

15. A genetic construct according to Claim 14 joined to a vector for introduction into an expression host, said vector comprising a marker for selection in said expression host.

16. A genetic construct according to Claim 15, wherein said host is a prokaryote.

17. A genetic construct according to Claim 15, wherein said host is a eukaryote.

18. A method for modulating homing of leukocytes to mucosal endothelial cells comprising MAdCAM-1, said method comprising:

adding to a cell population comprising leukocytes capable of binding to said MAdCAM-1 and endothelial cells expressing MAdCAM-1, a modulating amount of MAdCAM-1 or a fragment of MAdCAM-1 capable of interfering with the binding of said leukocyte with said endothelial cell.

19. A conjugate comprising MAdCAM-1 or fragment thereof of at least 12 amino acids covalently joined to a moiety capable of providing a detectable signal, an immunogen, or an agent capable of producing a physiological reaction.

20. A conjugate of at least 12 amino acid fragment of MAdCAM-1 and at least one of: a hapten; an antigen; a compound capable of providing a detectable signal; a toxin or an antibody; or a fused protein comprising said at least 12 amino acid fragment.

21. A method for screening for molecules capable of binding MAdCAM-1, said method comprising:

combining MAdCAM-1 or functional fragment thereof with said molecule; and

detecting the presence of complexes.

22. A method for screening for molecules capable of inhibiting binding of leukocytes to MAdCAM-1, said method comprising:

combining leukocytes with MAdCAM-1 and said molecule under conditions whereby leukocytes bind to MAdCAM-1; and

detecting the degree of binding of leukocytes with MAdCAM-1 as compared to the degree of binding of leukocytes to MAdCAM-1 in the absence of said molecule.

23. A method according to Claim 22, wherein said MAdCAM-1 is present in the cell membrane of a cell which does not normally express MAdCAM-1.

CCACGCGTCCGTCACAAGACAGAGGCAGGCATGGAATCCATCCTGGCCCTCCTGCTGGCCCT
GGCCCTAGTACCCTACCAGCTCAGCAGAGGACAGTCCTTCCAGGTGAACCCCCCTGAG
TCTGAGGTAGCTGTGGCCATGGGCACATCCCTCCAGATCACCTGCAGCATGTCCTGTG
ACGAGGGTGTAGCCCGGGTGCACTGGCGTGGTCTGGACACCAGCTTGGGAGTGTACAG
ACCCTCCCAGGCAGCAGTATCCTCTCTGTACGGGGCATGCTGTCAGACACAGGCACTC
CTGTGTGTGTGGGCTCCTGCGGGAGTCGAAGCTTCCAGCACTCCGTGAAGATCCTTGT
GTATGCCTTCCCAGACCAGCTGGTGGTGTCCCCGGAGTTCCTTGTACCTGGACAGGAC
CAGGTGGTGTCCTGCACGGCCCACAACATCTGGCCTGCAGACCCGAACAGTCTCTCCT
TTGCCCTGCTACTGGGAGAGCAGAGACTGGAGGGTGCCCAAGCCCTGGAACCAGAGCA
AGAAGAGGAGATACAAGAGGCTGAGGGCACACCACTGTTCCGAATGACACAACGCTGG
CGGTTACCCTCCCTGGGGACCCCTGCCCCTCCTGCCCTTCACTGCCAGGTCACCATGC
AGCTGCCCAAACTGGTGCTGACCCATAGAAAGGAGATTCCAGTACTACAGAGCCAGAC
CTCACCTAAGCCCCCCAACACGACCTCTGCTGAGCCCTACATCCTGACCTCATCAAGT
ACTGCTGAGGCAGTCTCCACTGGGCTCAACATCACCACCCTACCTTCTGCCCCTCCAT
ACCCCAAGCTTAGCCCTAGGACTCTGAGCTCTGAGGGACCTTGCCGCCCGAAAATCCA
CCAGGACCTGGAGGCAGGCTGGGAGCTACTCTGTGAAGCATCCTGTGGGCCCGGAGTT
ACTGTGCGCTGGACCTTGGCTCCTGGCGACCTGGCAACCTACCACAAGAGGGAGGCTG
GGGCCCAGGCATGGCTAAGCGTGCTGCCCCCAGGTCCCATGGTAGAGGGCTGGTTCCA
GTGCCGCCAGGACCCTGGCGGGGAGGTGACCAATCTGTATGTTCCTGGCCAGGTGACC
CCGAATTCCTCCTCCACCGTCGTCCTATGGATTGGCAGCTTGGTGCTGGGGCTGCTTG
CACTGGTCTTCCTTGCCTACCGCCTGTGGAAATGCTACCGGCCAGGTCCTCGCCCAGA
CACTAGCTCATGTACACACCTATGAAGCTCCATTATGCCAGACTAAAGGAGGCAGAGAGTGAC
CAGCTGCAGGATTTGGGGCATCAAGATGATAGTGTGGCCTCTTTCCTTGGTGGTCAGCACATCTAT
AAGTTTCTCCTGACTTCTGGGCTTTTCTGCCTGCTGGCCCAGAGCTAATAAAAGCCCCGTATCTA
AAAAAAAAAAAAAAA

MESILALLLALALVPYQLSRGQSFQVNPPESEVAVAMGTSLQITCSMSCDEGVARV
HWRGLDTSLGSVQTLPGSSILSVRGMLSDTGTPVCVGSCGSRSFQHSVKILVYAFPDQL
VVSPEFLVPGQDQVVSCTAHNIWPADPNSLSFALLLGEQRLEGAQALEPEQEEEIQEAE
GTPLFRMTQRWRLPSLGTPAPPALHCQVTMQLPKLVLTHRKEIPVLQSQTSPKPPNTT
SAEPYILTSSSTAEAVSTGLNITTLPSAPPYPKLSPRTLSSEGPCRPKIHQDLEAGWELL
CEASCGPGVTVRWTLAPGDLATYHKREAGAQAWLSVLPPGPMVEGWFQCRQDPGGQV
TNLYVPGQVTPNSSSTVVLWIGSLVLGLLALVFLAYRLWKCYRPGPRPDTSSCTH
L

FIGURE 1

INTERNATIONAL SEARCH REPORT

Int ional application No.
PCT/US93/11404

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/7.2, 240.2, 252.3, 320.1, 514/2; 530/350, 402, 866, 536/23.1, 24.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, BIOSIS, CHEM AB, EMBASE, DERWENT WPI, search terms: author names, addressin, madcam, mucosal, adhesion, leukocytes, meca-89, meca-367

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Science, Volume 233, Issued 01 August 1986, S. Jalkanen et al., "A distinct endothelial cell recognition system that controls lymphocyte traffic into inflamed synovium", pages 556-558, see entire document.	1-23
Y	Proceedings National Academy of Sciences, USA, Volume 88, Issued April 1991, J.V. Fecondo et al., "Inhibition of intercellular adhesion molecule 1-dependent biological activities by a synthetic peptide analog", pages 2879-2882, see entire document.	21-23

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 18 February 1994	Date of mailing of the international search report MAR 11 1994
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer RON SCHWADRON Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

Int. ional applicati n No.
PCT/US93/11404

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP, A2, 0,289,949 (SPRINGER et al.) 09 November 1988, see entire document.	1-6,10-20
Y	Nature, Volume 337, Issued 12 January 1989, M. Nakache et al., "The mucosal addressin is a tissue-specific endothelial cell adhesion molecule for circulating lymphocytes", pages 179-181, see entire document.	1-23

INTERNATIONAL SEARCH REPORT

Int tional application No.
PCT/US93/11404

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (5):

A61K 37/02; G01N 33/53; C07K 13/00; C07H 21/02, 21/04; C12N 5/12, 1/21

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/7.2, 240.2, 252.3, 320.1, 514/2; 530/350, 402, 866, 536/23.1, 24.1